

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-245042  
(43)Date of publication of application : 19.09.1997

(51)Int.Cl. G06F 17/30

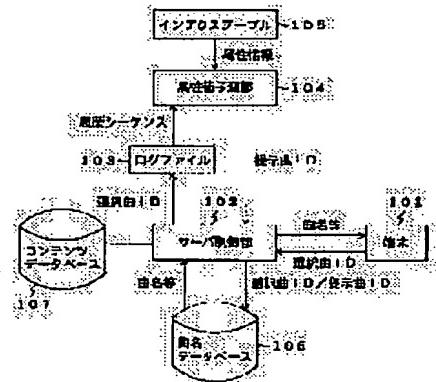
(21)Application number : 08-048735 (71)Applicant : NIPPON TELEGR & TELEPH CORP <NTT>  
(22)Date of filing : 06.03.1996 (72)Inventor : IHARA MASAYUKI  
KANEDA YOJI  
KANAYAMA HIDEAKI

## (54) RETRIEVAL SUPPORTING METHOD AND DEVICE THEREFOR

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To support the retrieval of a user by utilizing the attribute value sequence of the retrieval history information of the user and presenting information predicted to be desired in the next retrieval by the user.

**SOLUTION:** The history of the retrieval for which the user accesses a database 106 is held in a log file 103 as an information ID. An index table 105 manages attribute information provided in the information stored in the data base 106. An attribute value prediction part 104 uses the log file 103 and the index table 105, calculates an attribute value transition probability for each attribute of the retrieval history of the user and obtains the information ID corresponding to the attribute value in which the probability of appearing next to the attribute value transition of the information retrieved by the user lately is highest. A server control part 102 searches the information of the information ID from the data base 106 and sends it to a terminal 101.



### LEGAL STATUS

[Date of request for examination] 11.11.1998  
[Date of sending the examiner's decision of rejection]  
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]  
[Date of final disposal for application]  
[Patent number] 3212506  
[Date of registration] 19.07.2001  
[Number of appeal against examiner's decision of rejection]  
[Date of requesting appeal against examiner's decision of rejection]  
[Date of extinction of right]

THIS PAGE BLANK (USPTO)

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平9-245042

(43)公開日 平成9年(1997)9月19日

(51)Int.Cl.  
G 0 6 F 17/30

識別記号 庁内整理番号

F I  
G 0 6 F 15/403

技術表示箇所

3 4 0 B

審査請求 未請求 請求項の数2 O L (全5頁)

(21)出願番号 特願平8-48735

(22)出願日 平成8年(1996)3月6日

(71)出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72)発明者 井原 雅行

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(72)発明者 金田 洋二

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(72)発明者 金山 英明

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

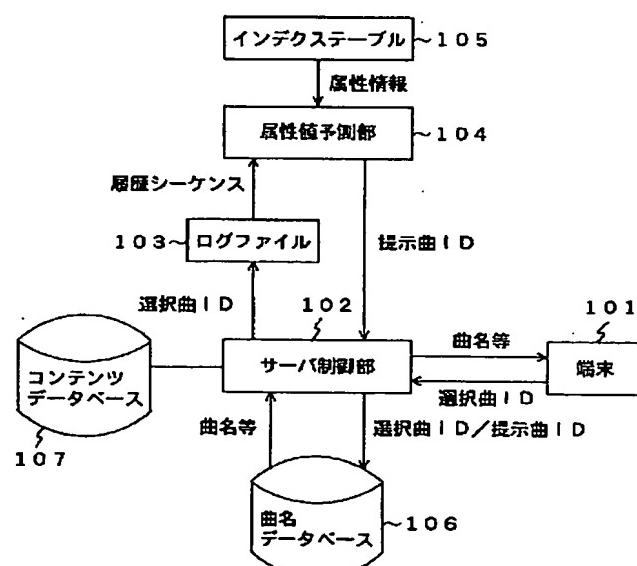
(74)代理人 弁理士 鈴木 誠

(54)【発明の名称】 検索支援方法および装置

(57)【要約】

【課題】 利用者の検索履歴情報の属性値シーケンスを活用して、利用者が次の検索で欲すると予想される情報を提示し、利用者の検索を支援する。

【解決手段】 利用者がデータベース106にアクセスした検索の履歴を情報IDとしてログファイル103として保持する。インデクステーブル105は、データベース106に蓄積されている情報が持つ属性情報を管理している。属性値予測部104は、ログファイル103とインデクステーブル105を用いて、利用者の検索履歴の各属性に対する属性値遷移確率を計算し、利用者の最近検索した情報の属性値遷移の次に出現する確率が最も高い属性値に対応する情報IDを得る。サーバ制御部102は、データベース106より該情報IDの情報を探しで端末101に送る。



**【特許請求の範囲】**

【請求項 1】 利用者がデータベースにアクセスして所望の情報を取り出す際に、次の検索でアクセスすると予想される情報を提示する検索支援方法であって、利用者が過去にデータベースにアクセスした情報の履歴を記憶しておき、利用者の検索履歴における連続する複数回の属性値遷移に関して、各情報の属性値から次に選択された情報の属性値への各属性毎の考え得る全ての遷移に対しての属性値遷移確率を計算し、該属性値遷移確率の中で、最も高い値である属性値間遷移に相当する情報を利用者に提示することを特徴とする検索支援方法。

【請求項 2】 利用者がデータベースにアクセスして所望の情報を取り出す際に、次の検索でアクセスすると予想される情報を提示する検索支援装置であって、利用者が過去にデータベースにアクセスした情報の履歴を記憶する履歴記憶手段と、データベースに蓄積されている情報が持つ属性情報を管理する属性情報管理手段と、前記履歴記憶手段と前記属性情報管理手段とを用いて、利用者の検索履歴における連続する複数回の属性値遷移に関して、各属性毎の考え得る全ての遷移に対する属性値遷移確率を計算する遷移確率計算手段とを備えることを特徴とする検索支援装置。

**【発明の詳細な説明】****【0001】**

【発明の属する技術分野】本発明は、利用者が音楽、映像等のデータベースにアクセスして所望の情報を取り出す情報検索システムにおいて、利用者が次の検索でアクセスすると予想される情報を、システム側で予測し、利用者に提示することにより、利用者の検索を支援する方法及び装置に関する。特に、データベースのリンクをたどって所望の情報を探すことの面倒さを解消しているという特徴と、利用者の検索履歴を連続的なシーケンスとして捉えているという特徴と、情報がもつ属性を用いて利用者に提示する情報を決定するという特徴をもつことにより、検索の専門家でない利用者が利用する情報検索システムに好適な検索支援方法及び装置に関するものである。

**【0002】**

【従来の技術】従来、一般に使用されている検索装置では、ツリー構造等のリンクをたどったり、あるいは、検索用のキーワードを入力することにより、利用者が欲しい情報を探している。しかし、検索の専門家でない利用者にとっては、深い階層構造や複雑なリンクをたどることは面倒であるという問題や、利用者が欲しい情報が曖昧である場合に満足できる情報を入手できないという問題がある。

【0003】これらの問題を解決する手法の一つとして検索履歴活用方法がある。これは、利用者の検索履歴を活用して、利用者が欲しいと予想される情報をあらかじめシステム側で予測し、利用者に提示するという方法で

ある。ところが、従来の検索履歴活用方法は、利用者がデータベース上のどの情報を何回アクセスしたかというアクセス回数を求めて、該アクセス回数が多い情報から優先順位を付けて提示するというものであった。

**【0004】**

【発明が解決しようとする課題】情報そのものに対するアクセス回数を用いる従来の検索履歴活用方法では、同じ情報が繰り返し利用者に提示されることがあるため、提示される情報に対する利用者の満足度が低くなるという問題がある。例えば、過去のアクセスにおける情報の属性値別のアクセス回数が最も多い属性値がZであったとする。従来の方法では、この属性値Zをもつ情報が利用者が欲しい情報であると予測して提示する。そして、最多アクセス回数の属性値がZ以外の属性値に変わるもので、属性値Zをもつ情報を提示し続けることになる。このため、利用者の好みが、ある属性値から別の属性値に変わっていく過渡期のケースや、一つの属性値に集中せずに複数の属性値を適度にアクセスするケース等では、確率の高い予測ができないという問題がある。

【0005】本発明は、従来の検索履歴活用方法の上記のような問題を解決すべくなされたものであり、その目的は、利用者の検索履歴を、情報そのものに対する単なるアクセス回数としてではなく、情報がもつ属性に関する連続性のある履歴シーケンスとして活用し、利用者が欲しいと予想される情報をあらかじめ予測してシステム側から提示して、その時々で変化する利用者の好みに合わせて提示する情報も変化させることにより、利用者の満足度を向上させることにある。

**【0006】**

【課題を解決するための手段】本発明では、利用者が最新アクセスした情報がもつ属性値を調べ、その属性値をもつ情報がアクセスされた次には、どの属性値をもつ情報がアクセスされやすいかという予測を、属性値の遷移パターンの遷移確率を用いて行う。具体的には、利用者がデータベースにアクセスした検索の履歴を情報のIDとして保持しているログファイルと、データベースに蓄積されている情報がもつ属性情報（例えば音楽ではジャンル、アーティストの性別等）を管理するインデクステーブルを持ち、このログファイルとインデクステーブルを用いて、利用者の検索履歴における連続する複数回の属性値遷移に関して、考え得る全ての属性値遷移パターンに対する属性値遷移確率を計算しておく。そして、利用者が最近アクセスした情報がもつ属性値から全属性値への遷移確率をそれぞれ調べ、最も遷移確率が高い属性値を予測属性値とする。こうして、各属性に対する予測属性値を求め、これらの予測属性値をもつ情報をインデクステーブルを参照してデータベースから探し、利用者に提示する。

**【0007】**

【発明の実施の形態】以下、図面によって本発明の一実

施例を説明する。なお、実施例では音楽情報の提示を題材にするが、勿論、本発明はこれに限定されるものでないことは云うまでもない。

【0008】図1は、本発明の一実施例を示す構成図である。図において、106は曲名やその他の所望情報を蓄積している曲名データベース、107は曲や映像情報等を蓄積しているコンテンツデータベースである。101は利用者が利用する端末、102は検索装置側にあって全体の制御を司る制御部（以下、サーバ制御部と称す）である。103は利用者がデータベース106、107にアクセスした検索の履歴を情報のID（ここでは曲のID）として保持するログファイル、104は利用者に提示する情報（ここでは曲名等）の属性値を予測する属性値予測部、105はデータベース106、107に蓄積されている情報がもつ属性情報（ここでは曲の属性：ジャンル、アーティストの性別等）を管理するインデクステーブルである。

【0009】利用者は端末101に表示された提示曲一覧から曲を選択する。曲が選択されると、選択曲のIDがサーバ制御部102に送られる。サーバ制御部102は、この選択曲IDを、利用者がデータベースにアクセスした検索の履歴を保持するログファイル103に格納する。ログファイル103には、利用者が曲選択を行う度に、選択曲IDが順次格納される。

【0010】属性値予測部104では、ログファイル103に格納されている選択曲IDの履歴シーケンスと、インデクステーブル105が管理している曲の属性情報をもとに、利用者に提示する曲がもつ属性値を予測し、対応する曲IDをサーバ制御部102に送る。サーバ制御部102は、当該曲IDをキーに曲名データベース106を参照し、曲名等を端末101に送る。

【0011】図2に、属性値予測部104の処理フローを示す。なお、図2は連続する2回の検索を考慮して提示曲IDを決定する場合の処理フロー例である。

【0012】ステップ201において、ログファイル103に格納されている選択曲IDによって作られる履歴シーケンスから、利用者が直前に選択した曲の曲IDを調べる。この選択曲IDに対して、ステップ202でインデクステーブル105を参照して、その選択曲がもつ属性値を調べる。曲にはジャンル、アーティストの性別といった複数の属性が考慮されているので、この選択曲にも複数の属性値が付与されている。ここで、直前の選択曲が持っている属性値を $A_s, B_s, C_s \dots$ とする。これらの属性値に対し、それぞれ、次の検索で選択されると予想される曲がもつ属性値 $A_g, B_g, C_g \dots$ を求める。ステップ203において、 $X = A$ を初期設定した後、ステップ204で、履歴シーケンスにおいて属性値 $X_s$  ( $X = A, B, C \dots$ ) の次に出現する確率が最も高い属性値 $X_g$ を探す。ステップ204→ステップ205→ステップ206→ステップ204…というループの繰

り返しにより、A, B, C…全ての属性について予測属性値 $X_g$ を求める。予測属性値 $X_g$ が求まったら、ステップ207において、インデクステーブル105を参照して予測属性値 $A_g, B_g, C_g, \dots$ をもつ曲のIDを探して、利用者への提示曲のIDとする。

【0013】以下に図3により、ステップ204の予測属性値を求める処理を、連続する2回の属性値遷移を考慮する例により示す。図3は連続する10回の検索結果における属性Aに関する属性値シーケンスである。なお、属性Aには、 $A_1, A_2, A_3$ という属性値が設定されているものとする。

【0014】いま、11回目の検索を予測することを考える。図3の属性値シーケンスにおいて、連続する2回の属性値遷移に着目すると、 $A_1$ から $A_1$ に遷移している場合は一回、 $A_1$ から $A_2$ に遷移している場合は三回、 $A_1$ から $A_3$ に遷移している場合は一回であり、その他に $A_2$ から $A_1$ に遷移している場合が三回、 $A_3$ から $A_1$ に遷移している場合が一回あり、属性値遷移は全部で九回ある。従って、各属性値遷移に対する遷移確率は、

$$A_1 \rightarrow A_1 : 1/9$$

$$A_1 \rightarrow A_2 : 3/9$$

$$A_1 \rightarrow A_3 : 1/9$$

$$A_2 \rightarrow A_1 : 3/9$$

$$A_3 \rightarrow A_1 : 1/9$$

となる。

【0015】ここでは、連続する2回の属性値遷移を考慮しているから、11回目の検索を予測するには、10回目の検索結果に着目して、10回目から11回目という遷移を考える。図3の属性値シーケンスによれば、10回目の検索における属性値は $A_1$ である。従って、属性値 $A_1$ からの属性値遷移確率を調べればよい。 $A_1$ からの遷移確率が最も高い属性値は $A_2$ であるから（遷移確率 $3/9$ ）、属性Aに関する求めるべき予測属性値（= $A_g$ ）は $A_2$ となる。その他の属性B, C, …についてもそれぞれ同様に、予測属性値 $B_g, C_g, \dots$ を求めることができる。また、連続する3回、あるいは3回以上の属性値遷移を考慮する場合も、同様にして予測が可能である。

【0016】このようにして求めた $A_g, B_g, C_g \dots$ をもつ情報を利用者に提示すればよい。即ち、属性値予測部104では、インデクステーブル105を参照して、これら予測属性値 $A_g, B_g, C_g \dots$ をもつ曲のIDを探し出し、それを提示曲IDとしてサーバ制御部102に送る。サーバ制御部102は、曲名データベース106を参照して、提示曲の曲名、アーティスト名等を調べ、端末101に提示曲一覧を表示する。そして、利用者が所望の曲名を指定すれば、サーバ制御部102は、コンテンツデータベース107を参照して、曲の内容等を端末101に送る。

【0017】

【発明の効果】以上のように、本発明は、利用者の検索履歴を、情報に対する単なるアクセス回数としてではなく、連続性のある履歴シーケンスとして活用し、利用者が欲しいと予想される情報をあらかじめシステム側で予測して利用者に提示し、利用者の検索を支援するものである。

【0018】本発明では、前回までの検索履歴をもとに、次にアクセスされると予想される情報がシステム側から提示されるため、利用者は、深い階層構造や複雑なリンクをたどるという面倒な操作や、あるいは、検索用のキーワードを入力するという操作なしで、欲しい情報を入手することができる。

【0019】しかも、本発明は、利用者の検索履歴を、情報そのものに対する単なるアクセス回数としてではなく、情報がもつ属性に着目して連続性をもった属性値シーケンスとして活用することにより、その時々で変化する利用者の好みに合わせて提示する情報も変化させるこ

とが可能である。そのため、情報そのものに対する単なるアクセス回数として予測を行う従来の方法に比べて、利用者の満足度を向上させることができる。

#### 【図面の簡単な説明】

【図1】音楽情報の提供を題材にした本発明の一実施例を示す構成図である。

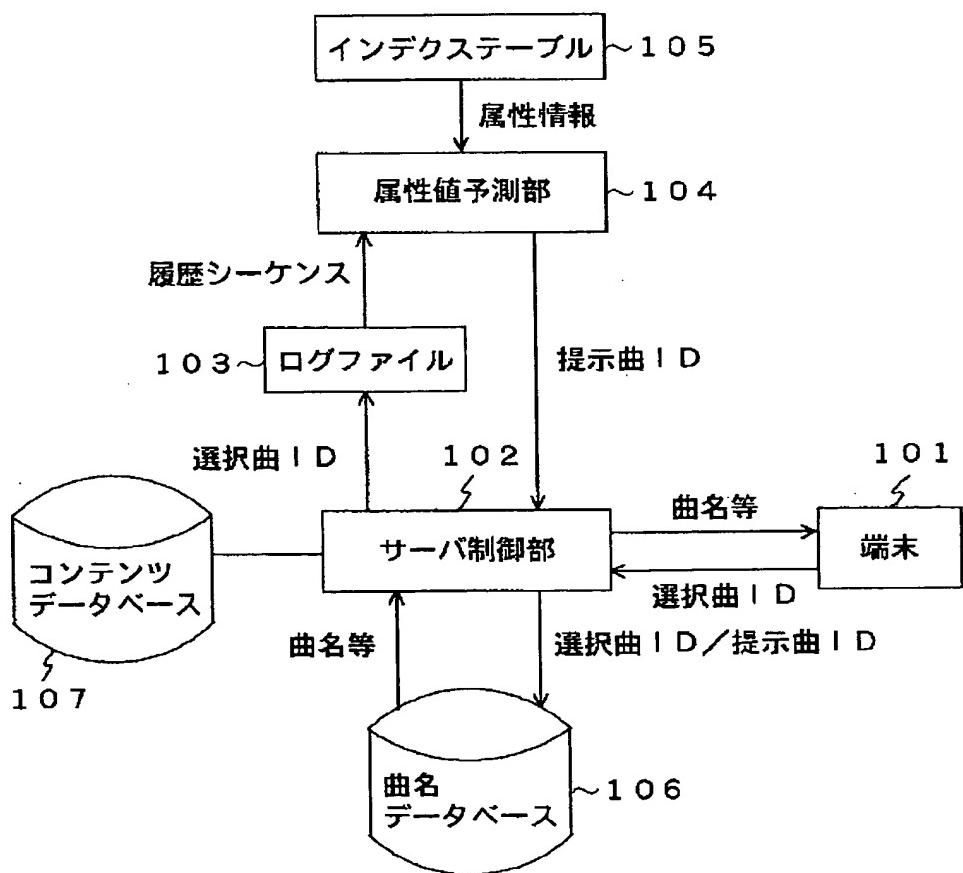
【図2】属性値予測部の処理フロー例を示す図である。

【図3】利用者の検索履歴における属性値シーケンスの一例である。

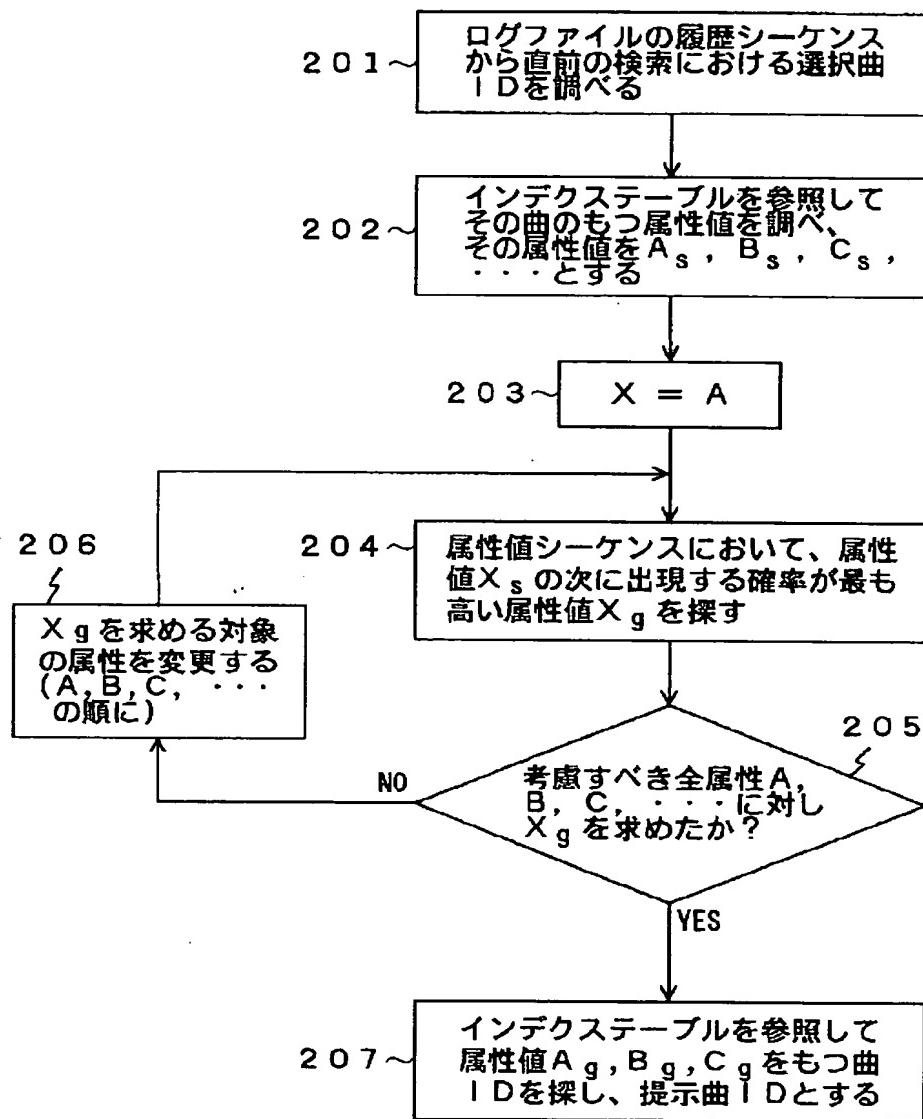
#### 【符号の説明】

- 101 端末
- 102 サーバ制御部
- 103 ログファイル
- 104 属性値予測部
- 105 インデクステーブル
- 106 曲名データベース
- 107 コンテンツデータベース

【図1】



【図2】



【図3】

検索項目	1	2	3	4	5	6	7	8	9	10	11
属性値	A <sub>1</sub>	A <sub>1</sub>	A <sub>2</sub>	A <sub>1</sub>	A <sub>3</sub>	A <sub>1</sub>	A <sub>2</sub>	A <sub>1</sub>	A <sub>2</sub>	A <sub>1</sub>	?

**THIS PAGE BLANK (USPTO)**